British Journal of Educational Technology (2013)
doi:10.1111/bjet.12129

The reverse modality effect: Examining student learning from interactive computer-based instruction

Fethi A. Inan, Steven M. Crooks, Jongpil Cheon, Fatih Ari, Raymond Flores, Murat Kurucay and Dmitrii Paniukov

Fethi A. Inan is an Associate Professor of Instructional Technology at Texas Tech University. His research interests are adaptive learning systems, individual differences, and technology integration. Steven M. Crooks is an Associate Professor of Instructional Technology at Texas Tech University. His research interests concern the discovery and application of research-based principles from the learning and instructional sciences to improve human learning. Jongpil Cheon is an assistant professor in the Instructional Technology program at Texas Tech University. His research interests involve implementing immersive online learning environments and investigating advanced technologies for interactive learning. Fatih Ari is a Programmer/Analyst in the Office of Institutional Planning and Assessment at Texas Tech University Health Sciences Center. His current research includes effective feedback design, design and development of adaptive learning environments, multimedia learning, and educational data mining. Raymond Flores is an assistant professor of Mathematics Education at Texas Tech University. His research interests include adaptive learning environments for mathematics instruction, pre-service mathematics teacher preparation, and technology integration in mathematics teaching. Murat Kurucay is a doctoral student in the Instructional Technology program at Texas Tech University. His main research interests are student self-regulation and motivation in online education, and multimedia learning. Dmitrii Paniukov is a doctoral student in the Applied Cognitive Psychology program at Texas Tech University. His main research interests are human learning, cognitive neuroscience and computational modeling of cognitive processes. Address for correspondence: Dr Fethi Inan, Educational Instructional Technology, College of Education, EDIT, Texas Tech University, Lubbock, TX 79409, USA. Email: fethi.inan@ttu.edu

Abstract
The purpose of this study was to explore the effects of modality on learning from multimedia instruction. This study utilized a factorial between-subject design to examine the effects of modality on student learning outcomes, study patterns and mental effort. An interactive computer-presented diagram was developed to teach the places of articulation in human speech. A total of 151 undergraduate students at a large southwestern university in USA participated in this study. Participants were randomly assigned to one of two modality conditions (ie, written text and spoken text). Data were obtained through surveys, student logs and knowledge tests. Findings revealed a reverse modality effect, wherein participants who studied with written text outperformed those who studied with spoken text.

Modality and the reverse modality effect

The modality design principle is used by instructional designers to foster the learning process in multimedia instruction. According to the modality principle, spoken text is more effective than written text when combined with pictures or animations (Mayer, 2009). Indeed, instruction incorporating the modality effect has been shown to improve performance on retention and transfer tests and to reduce mental effort (Brünken, Plass & Leutner, 2004; Mayer & Moreno, 2002; Tabbers, 2002). The theoretical rationale for the modality effect is based upon the notion that in working memory, spoken text is processed via the auditory channel while pictures and written text are processed via the visual channel (Mayer, 2008, 2009; Paivio, 1990). Therefore, when text and pictures are both processed via the visual channel of working memory, it becomes

Practitioner Notes

What is already known about this topic

- The modality principle suggests that spoken text is more effective than written text when combined with visuals.
- Recent research findings have questioned the modality principle suggesting certain learning conditions are favorable to a reverse modality effect in multimedia learning.

What this paper adds

- Investigated the boundaries of both modality and reverse modality effects.
- Examined the modality principle under instructional conditions that prior research suggests may be conducive to a reverse modality effect.
- Provided concrete recommendations for instructional designers.

Implications for practice and/or policy

- Written text was superior to spoken text when learners were given control over the pace and sequence of multimedia instruction.
- Written text is optimal during multimedia instruction when students are learning new and complex content because it allows them to employ text processing strategies.

overloaded. On the contrary, when spoken text is processed via the auditory channel and pictures via visual channel, no channel is overloaded. This arrangement fosters learning because spoken text and pictures do not compete for limited working memory resources.

The effectiveness of instruction with pictures and spoken text however may be limited to certain learning conditions such as when the instruction is complex and/or learning is self-paced (Cheon, Crooks, Inan, Flores & Ari, 2011; Tabbers, Martens & van Merriënboer, 2004). When studying relatively complex topics (eg, advanced mathematical equations, a second language or anatomy) or when learners have control of narration, the reverse modality effect may appear (Clark & Mayer, 2011). The reverse modality effect occurs when students have better learning outcomes when studying from instruction where text is presented visually rather than in an auditory format (Crooks, Cheon, Inan, Flores & Ari, 2012; Singh, Marcus & Ayres, 2012; Tabbers *et al*, 2004). These scholars associate the reverse modality effect to learners' ability to study at their own pace and the use of text processing strategies to understand complex written text (Byrne & Curtis, 2000; Corston & Colman, 1997; Stiller, Freitag, Zinnbauer & Freitag, 2009; Tabbers & de Koejier, 2010).

Although the number of computer and web-based multimedia instruction has increased drastically, specific guidelines for how to design multimedia instructions are not entirely clear in regard to modality (Ginns, 2005) and, therefore, additional research is needed to clarify in which learning context instructional designer should consider the modality or reverse modality effect. Although the modality effect has received considerable empirical support (Brünken *et al*, 2004; Mayer, 2009; Mayer & Moreno, 2002), some recent experiments have found the opposite to modality—the reverse modality effect (Cheon *et al*, 2011; Crooks *et al*, 2012). Therefore, practitioners continue to be left undirected and have no specific prescriptions of which modality design principle is more suitable in diverse instructional conditions. This study aimed to further investigate the boundaries of both modality and reverse modality effects, providing concrete recommendations for instructional designers.

Purpose of the study

The purpose of this study was to investigate the modality/reverse modality effect when learning from a computer-based interactive diagram. The research question explored was "Do the written

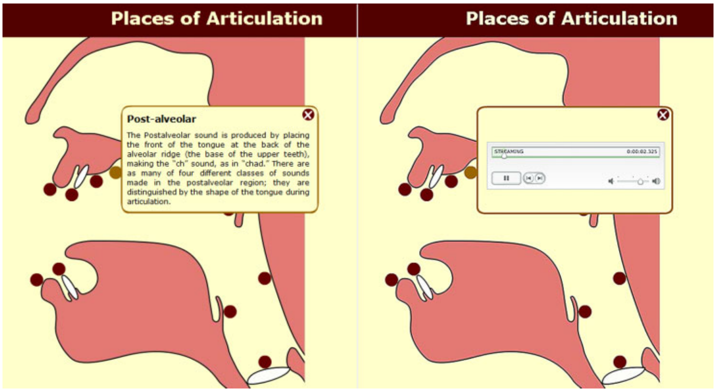

Figure 1: Modality conditions

text and the spoken text groups differ in learning performance, study patterns and mental effort?"

Method

Participants

One hundred and fifty-one students enrolled in an undergraduate course on computer literacy from a large southwestern university in USA participated in this study. Participants were randomly assigned to one of the two experimental conditions: written text group ($n = 79$) and spoken text group ($n = 72$). Participants' ages ranged from 17 to 37 years but most (95%) were younger than 23 years. The majority of the students were white (69.5%) and about two-thirds of them were female (64.9%).

Instructional material

The participants studied the 12 places of articulation in human speech via a computer-presented diagram of a cross-section of the human head. Each place of articulation was accompanied by descriptions explaining the speech sounds formed at that location. Two versions of the computer diagram were created: (1) a written text condition that provided on-screen text descriptions and (2) a spoken text condition that provided audio descriptions (see Figure 1). Dependent on their condition, once students clicked a location marker, they would either see a pop-up with written text or listen to the same text in spoken form. In the spoken text condition, students were also able to increase/decrease volume, pause, rewind and fast forward through the audio by clicking respective buttons in the pop-up. Participants were given 10 minutes to study the instruction having complete control over the pace and sequence of their learning. After study time, participants were asked to complete a mental effort measure and a series of knowledge assessment tests.

Data collection instruments

Knowledge assessments

In order to examine the impact of the modality condition on student learning outcomes, various assessments were used in this study (eg, a comprehension test, a matching test, a spatial labeling test and a spatial reconstruction test) to measure the participants' performance after studying the instructional material. A team of researchers and graduate students worked on the instruments to ensure that they were aligned with the research questions and instructional content. Data collection procedures and instrumentation built on those provided by previous studies (Cheon *et al*, 2011; Crooks *et al*, 2012; Tabbers *et al*, 2004).

The comprehension test consisted of a 12-item multiple choice instrument designed to measure student comprehension of the content presented in the computer program. The matching test consisted of a listing of the 12 places of articulation (eg, dental) and a corresponding list of the sounds made at each place (eg, "th" as in thunder), plus three distracters. Participants were required to match each place of articulation with its corresponding sound in order to assess how well they associated the place of articulation with a specific sound it makes. The spatial labeling test provided unlabeled articulation places in the diagram and the list so that participants can move a description to the right place. This test score was calculated by summing up the number of items correctly placed. Finally, the reconstruction test presented each participant with a blank outline of the places of articulation diagram and a listing of the 12 places of articulation next to the diagram. Each participant was required to drag the name of each place to its correct location on the diagram. The reconstruction test score reflect the distance between correct locations and the locations participants moved.

Mental effort

A commonly used subjective measure of mental effort scale was administered to measure the amount of cognitive resources that participants' perceived that they invested while studying (Kalyuga & Sweller, 2005; Paas, Tuovinen, Tabbers & van Gerven, 2003). This scale consisted of a 9-point Likert scale ranged from 1 (*extremely low*) to 9 (*extremely high*). Paas and van Merrienboer (1994) evaluated the scale as a reliable instrument ($\alpha = .82$).

Study patterns

Data logs were kept to track participants' actions (eg, total number of clicks and study time) while they were interacting with the instructional material. The first, number of clicks, counted the number of times that each participant clicked on place of articulation markers to read or hear the description of each place of articulation. The second, study time, tracked the total number of seconds each participant spent reviewing or listening to descriptions.

Results

Equivalence of groups

Although random assignment was used to account for the differences between the treatment and control groups, a preliminary data analysis was conducted to examine if there were any significant variations in the group assignments. The percentage of participants in each group did not differ in linguistics background (χ^2 (1, $n = 151$) = .013, $p = .908$), gender (χ^2 (1, $n = 151$) = 1.62, $p = .203$), ethnicity (χ^2 (4, $n = 151$) = 4.87, $p = .301$), learning styles (χ^2 (1, $n = 151$) = 2.66, $p = .606$) and academic classification (χ^2 (3, $n = 151$) = 4.94, $p = .176$). Similarly, the result of an independent t-tests yielded no significant differences between conditions in regard to age ($t(149) = 1.31$, $p = .191$). Additionally, participants' familiarity with the content of the instructional material (linguistics) was collected prior to study and the data were examined with and without those 13 participants who reported familiarity with the content. There were no meaningful changes in the results when those participants were excluded from the analysis. Considering the random assignment, equal distributions of participants with linguistics background in each group and no-significant differences in findings with the exclusion of these participants, all participants were included in the reporting findings.

Group differences

In order to answer research questions, a series of factorial data analysis was used. Multivariate analysis of variance (MANOVA) was used to examine differences between the modality conditions (written text vs. spoken text) on comprehension, matching, spatial reconstruction and spatial labeling. Further, separate independent sample t-tests were conducted to investigate the differences between treatment groups on number of clicks, study time and mental effort. Effect sizes (ES) were computed using Cohen's d formula (Cohen, 1988) to determine the educational importance of differences. Cohen suggested that .2 be considered a small ES, .5 represents a medium ES and .8 a large ES.

Table 1: The means, standard deviations, and results of inferential test scores

Measures		*n*	*M*	*SD*	*t*	*p*	*ES*
Matching	Written text	79	3.76	2.77	2.68	.008*	.44
	Spoken text	72	2.65	2.25			
Comprehension	Written text	79	7.61	2.98	2.50	.013*	.41
	Spoken text	72	6.40	2.92			
Reconstruction†	Written text	79	37.29	32.16	−6.50	.000*	1.05
	Spoken text	72	87.67	60.10			
Labeling	Written text	79	10.28	2.78	4.70	.000*	.76
	Spoken text	72	7.72	3.86			
Study time	Written text	79	529.64	168.05	−2.32	.022*	.38
	Spoken text	72	595.35	180.24			
# of clicks	Written text	79	51.58	22.62	10.56	.000*	1.75
	Spoken text	72	22.13	7.26			
Mental effort	Written text	79	6.46	1.82	.63	.533^{ns}	NA
	Spoken text	72	6.26	1.94			

Notes:
†Lower scores represent more accurate placement on reconstruction test.
*$p < .05$.
ES, effect size; M, mean; NA, not applicable; SD, standard deviation.

Firstly, MANOVA yielded significant multivariate effects for modality (*Wilks* $\lambda = .769$, $F(4,146) = 10.976$, $p < .001$). As a follow-up, separate independent *t*-tests were used to determine which of the four dependent variables significantly contributed to group differences. For all four dependent variables, participants in the written text group outperformed participants in the spoken text group. Results indicated reverse modality effects for matching $t(149) = 2.68$, $p < .001$, ES = .44; comprehension, $t(149) = 2.50$, $p < .001$, ES = .41; spatial reconstruction, $t(149) = -6.50$, $p < .001$, ES = 1.05; and spatial labeling, $t(149) = 4.70$, $p < .001$, ES = .76. Table 1 shows the means and standard deviations for each of the learning outcome measures by modality condition.

In regard to study patterns, results from an independent sample *t*-test revealed that participants in the written text group clicked on markers significantly more than the participants in the spoken text group, $t(149) = 10.56$, $p < .001$, ES = 1.75. On the other hand, findings revealed that the participants in the written text group spent significantly less time on reviewing marker descriptions than the participants in the spoken text group, $t(149) = -2.32$, $p < .05$, ES = .38.

Finally, results from an independent sample *t*-test yielded no significant modality effect on mental effort, $t(149) = .63$, $p = .533$. Although nonsignificant, participants in the written text group reported higher mental effort ratings than the participants in the spoken text group (see Table 1).

Discussions and conclusion

This study examined the effect of modality on student learning with multimedia instruction. Results indicated that learning improved when spoken text is replaced with written text when presented with a visual. The finding is supported by previous research, which similarly suggested that the reverse modality effect would be observed under certain instructional conditions (Cheon *et al*, 2011; Crooks *et al*, 2012; Tabbers *et al*, 2004). Instructional situations where spatial learning is an important learning outcome may be one of the reasons for observing the reverse modality effect (Cheon *et al*, 2011; Crooks *et al*, 2012). Furthermore, previous research suggests that the reverse modality effect may be observed in situations where spoken text contains complex information that is new to students (Clark & Mayer, 2011; Leahy & Sweller, 2011). In this study, participants studied topics with several linguistics terms that they were not familiar with.

Therefore, it can be suggested that designers should consider using written text when designing multimedia instruction with similar learning topics.

The modality effect is mainly observed in multimedia learning when learners have little or no control over the pacing and sequence of the instruction (Brünken *et al*, 2004; Mayer, 2009; Wong, Leahy, Marcus & Sweller, 2012). In the current study, learners had control over the pacing and sequence of instruction where they could revisit the content to understand it. The findings indicated that learners in the written text group clicked on each marker more frequently than the learners in the spoken text group. This may be due to the fact that learners in the written text group had more flexibility to visit various content elements, employ various text processing strategies and make more efficient use of cognitive resources (Byrne & Curtis, 2000; Furnham, de Siena & Gunter, 2002; Witteman & Segers, 2009). The findings are consistent with previous research, which suggests that written text may be more appropriate when learners are given more control over the instruction (Cheon *et al*, 2011; Crooks *et al*, 2012; Singh *et al*, 2012; Tabbers *et al*, 2004).

In regard to mental effort, results indicated no significant differences between the two modality conditions. Although nonsignificant, written text group reported slightly higher mental effort ratings than the spoken text group. The mental effort survey used in this study was a subjective, single-item scale. Subjective measures have been extensively used by cognitive load researchers and found to be reliable (Chen, Epps & Chen, 2011; Paas *et al*, 2003). However, subjective measures of mental effort have been criticized lately by some cognitive load researchers because it is unclear how subjective perceptions of mental effort are related to actual cognitive load (Brünken, Plass & Leutner, 2003). Additionally, Ayres (2006) pointed out that subjective measures have been mainly used for reflecting the total amount of cognitive load rather than individual ones. Therefore, subjective measures may not be sensitive enough to distinguish different types of cognitive load caused by different cognitive processes (Plass, Moreno & Brünken, 2010; Van Gog & Paas, 2008). Furthermore, subjective measures are relative and dependent upon learners' abilities to reflect on their cognitive processes and assess the mental effort invested during the learning process (Ayres, 2006). Hence, it might not be entirely clear which type of cognitive load learners refer to when answering a single-item mental effort measure. Leahy and Sweller (2011), for instance, asserted that if information is complex or unfamiliar, spoken text requires more cognitive resources because learners direct their attention to transient spoken text. This may have caused higher extraneous cognitive load imposed on learners' cognitive systems, thus a decrease in learning performance because learners had to allocate their cognitive resources for mainly holding information in memory. On the other hand, Cheon *et al* (2011) found that learners studying written text may direct more cognitive resources toward text processing strategies, and this may have caused higher germane cognitive load imposed on learners' cognitive systems, thus an increase in learning performance.

Being able to measure different types of cognitive load may help researchers when making inferences about increases or decreases in learning performance due to cognitive load. In addition to subjective effort ratings used in this study, future researchers should explore dual-task performance techniques, which use the performance of a task that is performed concurrently with a primary task (Brünken, Plass & Leutner, 2003, 2004) and psychophysiological measures such as eye movements (Amadieu, van Gog, Paas, Tricot & Marine, 2009), electroencephalograph (Antonenko, Paas, Grabner & van Gog, 2010) or neuroimaging (Whelan, 2007). These instruments and/or methods (eg, dual task, eye tracking) can be used to gauge the different types of cognitive load (Cheon & Grant, 2012; Leppink, Paas, Van der Vleuten, Van Gog & Van Merriënboer, 2013; Plass *et al*, 2010; Schmidt-Weigand, Kohnert & Glowalla, 2010).

Although there is still much to learn about the boundaries of the modality effect in multimedia learning, the current study's findings shed light on the effect of reverse modality principle in a learner-paced, complex multimedia learning environment. The results suggest that when multimedia is designed for complex content in a learned paced format, instructional developers should consider applying the reverse modality principle to promote student learning. However, the results should be interpreted with caution as their generalizability may be limited. First, the geographical area was limited, and participants were recruited from only one university. Second, all study participants were college students with limited age variations. Future research could include a replication of the study with the inclusion of multiple institutions and the solicitation of students from diverse age groups. There were other limitations of the current study that participants were exposed to the treatment materials for a relatively short amount of time and their performance was only measured immediately after the experiment. Future studies should re-examine current findings by employing instructional material with longer study time and testing student performance also under delayed conditions. One of the feasible approaches for future studies is to extend this study with inclusion of additional data such as collecting eye-tracking data, which could be particularly useful in generating explanations how reverse modality effect come about and how learners process instructional materials (Schmidt-Weigand, Kohnert & Glowalla, 2010; Van Gog & Scheiter, 2010).

References

Amadieu, F., van Gog, T., Paas, F., Tricot, A. & Marine, C. (2009). Effects of prior knowledge and concept-map structure on disorientation, cognitive load, and learning. *Learning and Instruction*, *19*, 5, 376–386.

Antonenko, P., Paas, F., Grabner, R. & van Gog, T. (2010). Using electroencephalography to measure cognitive load. *Educational Psychology Review*, *22*, 4, 425–438.

Ayres, P. (2006). Using subjective measures to detect variations of intrinsic cognitive load within problems. *Learning and Instruction*, *16*, 5, 389–400.

Brünken, R., Plass, J. & Leutner, D. (2003). Direct measurement of cognitive load in multimedia learning. *Educational Psychologist*, *38*, 53–61.

Brünken, R., Plass, J. L. & Leutner, D. (2004). Assessment of cognitive load in multimedia learning with dual-task methodology: auditory load and modality effects. *Instructional Science*, *32*, 115–132.

Byrne, M. & Curtis, R. (2000). Designing health communication: testing the explanations for the impact of communication medium on effectiveness. *British Journal of Health Psychology*, *5*, 189–199.

Chen, S., Epps, J. & Chen, F. (2011). A comparison of four methods for cognitive load measurement. In *Proceedings of the 23rd Australian Computer-Human Interaction Conference* (pp. 76–79). ACM.

Cheon, J., Crooks, S., Inan, F., Flores, R. & Ari, F. (2011). Exploring the instructional conditions for a reverse modality effect in multimedia instruction. *Journal of Educational Multimedia and Hypermedia*, *20*, 2, 117–133.

Cheon, J. & Grant, M. (2012). Examining the relationships of different cognitive load types related to user interface in web-based instruction. *Journal of Interactive Learning Research*, *23*, 1, 29–55.

Clark, R. C. & Mayer, R. E. (2011). *E-learning and the science of instruction: proven guidelines for consumers and designers of multimedia learning*. San Francisco, CA: Pfeiffer.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Corston, R. & Colman, A. M. (1997). Modality of communication and recall of health-related information. *Journal of Health Psychology*, *2*, 2, 185–194.

Crooks, S., Cheon, J., Inan, F., Flores, R. & Ari, F. (2012). Modality and cueing in multimedia learning: examining cognitive and perceptual explanations for the modality effect. *Computers in Human Behavior*, *28*, 3, 1063–1071.

Furnham, A., de Siena, S. & Gunter, B. (2002). Children's and adults recall of children's news stories in both print and audio-visual presentation modalities. *Applied Cognitive Psychology*, *16*, 2, 191–210.

Ginns, P. (2005). Meta-analysis of the modality effect. *Learning and Instruction*, *14*, 4, 313–331.

Kalyuga, S. & Sweller, J. (2005). Rapid dynamic assessment of expertise to improve the efficiency of adaptive e-learning. *Educational Technology Research and Development*, *53*, 3, 83–93.

Leahy, W. & Sweller, J. (2011). Cognitive load theory, modality of presentation and the transient information effect. *Applied Cognitive Psychology*, *25*, 943–951.

Leppink, J., Paas, F., Van der Vleuten, C. P., Van Gog, T. & Van Merriënboer, J. J. (2013). Development of an instrument for measuring different types of cognitive load. *Behavior Research Methods*, *45*, 1–15.

Mayer, R. E. (2008). *Learning and instruction* (2nd ed.). Upper Saddle River, NJ: Pearson Merrill Prentice Hall.

Mayer, R. E. (2009). *Multimedia learning* (2nd ed.). New York: Cambridge University Press.

Mayer, R. E. & Moreno, R. (2002). Animation as an aid to multimedia learning. *Educational Psychology Review*, *14*, 1, 87–99.

Paas, F., Tuovinen, J. E., Tabbers, H. & van Gerven, P. W. M. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist*, *38*, 63–72.

Paas, F. & van Merrienboer, J. (1994). Variability of worked examples and transfer of geometrical problem-solving skills: a cognitive-load approach. *Journal of Educational Psychology*, *86*, 1, 22–133.

Paivio, A. (1990). *Mental representations: a dual coding approach*. New York: Oxford University Press.

Plass, J. L., Moreno, R. & Brünken, R. (2010). *Cognitive load theory*. New York: Cambridge University Press.

Schmidt-Weigand, F., Kohnert, A. & Glowalla, U. (2010). Explaining the modality and contiguity effects: new insight from investigating students' viewing behavior. *Applied Cognitive Psychology*, *24*, 2, 226–237.

Singh, A. M., Marcus, N. & Ayres, P. (2012). The transient information effect: investigating the impact of segmentation on spoken and written text. *Applied Cognitive Psychology*, *26*, 6, 848–853.

Stiller, K. D., Freitag, A., Zinnbauer, P. & Freitag, C. (2009). How pacing of multimedia instructions can influence modality effects: a case of superiority of visual texts. *Australian Journal of Educational Technology*, *25*, 2, 184–203.

Tabbers, H. B. & de Koejier, B. (2010). Learner control in animated multimedia instructions. *Instructional Science*, *38*, 441–453.

Tabbers, H. K. (2002). *The modality of text in multimedia instructions. Refining the design guidelines*. Heerlen: Open University of the Netherlands.

Tabbers, H. K., Martens, R. L. & van Merriënboer, J. J. G. (2004). Multimedia instructions and cognitive load theory: effects of modality and cueing. *British Journal of Educational Psychology*, *74*, 71–81.

Van Gog, T. & Paas, F. (2008). Instructional efficiency: revisiting the original construct in educational research. *Educational Psychologist*, *43*, 1, 16–26.

Van Gog, T. & Scheiter, K. (2010). Eye tracking as a tool to study and enhance multimedia learning. *Learning and Instruction*, *20*, 2, 95–99.

Whelan, R. R. (2007). Neuroimaging of cognitive load in instructional multimedia. *Educational Research Review*, *2*, 1, 1–13.

Witteman, M. J. & Segers, E. (2009). The modality effect tested in children in a user-paced multimedia environment. *Journal of Computer Assisted Learning*, *26*, 2, 132–142.

Wong, A., Leahy, W., Marcus, N. & Sweller, J. (2012). Cognitive load theory, the transient information effect and e-learning. *Learning and Instruction*, *22*, 6, 449–457.